5a

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER OF PATENTS AND TRADEMARKS
Washington, D.C. 20231
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/263,068	03/08/1999	BORIS PECHENY	50277-164	1360

7590 05/31/2002

DITTHAVONG & CARLSON, P.C.
10507 BRADDOCK RD
SUITE A
FAIRFAX,, VA 22032

EXAMINER	
FLEURANTIN, JEAN B	
ART UNIT	**PAPER NUMBER**
2172	

DATE MAILED: 05/31/2002

Please find below and/or attached an Office communication concerning this application or proceeding.

PTO-90C (Rev. 07-01)

Office Action Summary	Application No.	Applicant(s)
	09/263,068	PECHENY, BORIS
	Examiner	**Art Unit**
	Jean B Fleurantin	2172

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133).
- Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *Consideration 05/18/2002* .

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1-34* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1-34* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

11)☐ The proposed drawing correction filed on _____ is: a)☐ approved b)☐ disapproved by the Examiner.

If approved, corrected drawings are required in reply to this Office action.

12)☐ The oath or declaration is objected to by the Examiner.

Priority under 35 U.S.C. §§ 119 and 120

13)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____ .

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

14)☐ Acknowledgment is made of a claim for domestic priority under 35 U.S.C. § 119(e) (to a provisional application).

a) ☐ The translation of the foreign language provisional application has been received.

15)☐ Acknowledgment is made of a claim for domestic priority under 35 U.S.C. §§ 120 and/or 121.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☐ Information Disclosure Statement(s) (PTO-1449) Paper No(s) _____ .
4) ☐ Interview Summary (PTO-413) Paper No(s). _____ .
5) ☐ Notice of Informal Patent Application (PTO-152)
6) ☐ Other: .

DETAILED ACTION

Response to Amendment

1. Claims 1-34 are remained for examination.

2. Applicant's request for reconsideration of the finality of the rejection of the last Office action mailed on 11/23/2001 is persuasive and, therefore, the finality of that action is withdrawn.

Response to Applicant' Remarks

3. Applicant stated on pages 2-3 that the rejection is respectfully traversed because Li et al., alone or in combination with Levine et al., fails to disclose, teach, or otherwise suggest the limitations of the claims. For example, the claims require either "identifying a lexical container from among a plurality of lexical containers based on **a length of a key**" or "identifying a hash table from among a plurality of hash tables based on **a length of the key**". However, Examiner disagrees because Cohen includes steps of calculating m delaying additively complementary to the m n-gram selection positions respectively for the distinct length of the keywords in the dictionary; which is read as a length of the key (see col. 4, lines 50-53); also in column 6, lines 50 through 52 Cohen further teaches steps of one may record the lengths of the keywords associated with each hash address produced by H_q. Thus, it would have been obvious to a person of ordinary skill in the art at the time the invention was made to modify the teachings of Li and Cohen with the step of a length of the key. This modification would allow the teachings of Li and Cohen to improve the accuracy and the reliability of the lexical cache.

Examiner is entitled to give claim limitations their broadest reasonable interpretation in light of the specification.

Interpretation of Claims-Broadest Reasonable Interpretation

During patent examination, the pending claims must be 'given the broadest reasonable interpretation consistent with the specification.' Applicant always has the opportunity to amend the claims during prosecussion and broad interpretation by the examiner reduces the possibility that the claim, once issued, will be interpreted more broadly than is justified. In re Prater, 162 USPQ 541,550-51 (CCPA 1969).

Claim Rejections - 35 U.S.C. § 103

4. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

Claims 1-34 are rejected under 35 U.S.C. 103(a) as being unpatentable over Li et al. (US Pat. No. 5,774,588) in view of Cohen (US Pat. No. 6,169,969) ("Li"), ("Cohen").

As per claims 1, 16 and 31-32 Li teaches a method of searching for a string in a lexical cache (thus, direct comparison of an input string to a large number of dictionary entries comparing n-gram representations may also consume a large amount of computational time, also if a system encodes an entire lexicon with non-positional n-gram encoding a match does not mean that the string is a dictionary word this ambiguity arises because one only determines if there is a

common characteristic with the dictionary as a whole, on the other hand n-gram comparisons have the benefit of simple binary inexact matching are faster than many other comparison schemes and can save some space when the string and dictionary entries are hashed to sets of possible n-grams; which is readable as searching for a string in a lexical cache) (see col. 2, lines 15-26), as claimed comprises the computer implemented steps of generating a key based on the string (thus, this string is assumed to be the 45th entry in the lexicon, which is equivalent to generating a key based on the string) (see figure 5, col. 8, lines 28-31);

searching the lexical container for an entry associated with the string based on the key (thus, direct comparison of an input string to a large number of dictionary entries comparing n-gram representations may also consume a large amount of computational time, also if a system encodes an entire lexicon with non-positional n-gram encoding, a match does not mean that the string is a dictionary word this ambiguity arises because one only determines if there is a common characteristic with the dictionary as a whole, on the other hand n-gram comparisons have the benefit of simple binary inexact matching are faster than many other comparison schemes and can save some space when the string and dictionary entries are hashed to sets of possible n-grams; which is readable as searching the lexical container for an entry associated with the string) (see col. 2, lines 15-26). But, Li does not explicitly indicate the step of ~~the~~ a length of the key. However, implicitly Cohen strongly teaches step of calculating m delaying additively complementary to the m n-gram selection positions respectively for the distinct length of the keywords in the dictionary; which is readable as a length of the key (see col. 4, lines 50-53). Also

in column 6, lines 50 through 52 Cohen further teaches steps of one may record the lengths of the keywords associated with each hash address produced by H_q. Thus, it would have been obvious to a person of ordinary skill in the art at the time the invention was made to modify the teachings of Li and Cohen with the step of a length of the key. This modification would allow the teachings of Li and Cohen to improve the accuracy and the reliability of the lexical cache, and provide a method for full text dictionary string matching using n gram hashing (see col. 3, lines 35-37).

As per claims 2, 4, 17 and 19, Li teaches a method as claimed, wherein the step of generating a key based on the string includes the step of compressing the string to produce the key (thus, indexing of the lexicon entries produces a fixed result for a given lexicon, which is readable as generating a key based on the string includes the step of compressing the string to produce the key) (see col. 3, lines 27-28).

As per claims 3 and 18, Li teaches a method as claimed, wherein the step of compressing the string to produce the key includes the step of performing an n-gram compression on the string (thus, the step of partitioning a representation of an entry comprises forming an n-gram vector representing the entry folding the n-gram vector into a signature vector of reduced bit length; which is readable as wherein the step of compressing the string to produce the key includes the step of performing an n-gram compression on the string) (see col. 3, lines 1-4); also in column 14, lines 49 through 50, Li further teaches steps of folding said n-gram vector into a signature vector by combining multiple n-gram into bits.

As per claims 5 and 20, the limitations of claims 5 and 20 are rejected in the analysis of claim 1 above, and these claims are rejected on that basis.

As per claims 6 and 21, in addition to the discussion in claim 1 above, Li further teaches each of said sequences of slots corresponding to a respective hash value (thus, partitioning the signature vector into groups each having a predetermined number of bits, the bits of the groups are preferably arranged in descending order of frequency of appearance of each bit in the lexicon; which is readable as sequences of slots corresponding to a respective hash value) (see col. 3, lines 5-8);

computing a hash value based on the key (thus, the method begins with the partitioning and hashing step which is the least computationally expensive it then proceeds to the encoded vector comparison step and finally the direct string comparison of the edit distance step which is the most computationally expensive; which is readable as computing a hash value based on the key) (see col. 14, lines 19-24).

As per claims 7 and 22, Li substantially teaches a method as claimed, wherein the step of computing a hash value based on the key includes the step of computing the hash value based on the key and a prime number associated with the hash table (thus, the partitioning and hashing subroutine of figure 2 is run for each entry of the lexicon the indexing or hashing step 240 is carried out for all the entries of the lexicon when the entire lexicon is processed, many of the individual bucket addresses will be associated with varying numbers of the entries of the lexicon all the foregoing steps are completed in advance of actual matching to any unverified string the

group partitioning and the linked list table 40 are fixed for a particular lexicon; which is readable as computing a hash value based on the key includes the step of computing the hash value based on the key and a prime number associated with the hash table) (see col. 8, lines 19-27).

As per claims 8 and 23, in addition to the discussion in claims 1 and 6 above, Li teaches wherein the step of searching the hash table based on the hash value includes the steps of indexing one or more fixed regions of the hash table (thus, method for comparing unverified strings to the entries of a dictionary or lexicon in order to reduce the valid dictionary candidates to be considered as possible correct matches for the unverified string, there is also a need for such a method which does not exclude relevant parts of the lexicon and is capable of generating a short list of candidates that have a high likelihood of including an accurate match; which is readable as searching the hash table based on the hash value includes the steps of indexing one or more fixed regions of the hash table) (see col. 2, lines 34-40).

As per claims 9 and 24, Li substantially teaches a method as claimed, wherein the step of searching the hash table further includes the step of searching for the key in a linked list of slots stored in an expansion region of the hash table, if the key was not found in the one or more respective slots for the key; these numbers are used to create a bucket address table 40 each bucket address may eventually have more than one bucket as shown at address 0 if needed to accommodate multiple signature vectors indexed to that address the capacity of each bucket may be, for example 100 pointers all the buckets sharing a bucket address form a linked list; which is readable as searching for the key in a linked list of slots stored in an expansion region of the hash

table) (see col. 7, lines 54-60); also in column 7, lines 63 through 6 Li further teaches where these decimal numbers for each of the seven groups of the partitioned signature vector 30-36 are used to hash the signature vector to up seven linked lists.

As per claims 10 and 25, in addition to the discussion in claim 9 above, Li further teaches then moving a relative position of the entry for the string within the sequence of slots toward the beginning of the sequence of slots (thus, partition the signature vector of a particular lexicon entry the same bits are compared to the bit mask only if the bit is set to '1' in the entry's signature vector is the corresponding bit of the first group set to '1' for example if bit no. 2 'AI through AP' contains the bi-grams having the most frequent occurrence in the lexicon bit $g_{1,1}$ will be given a value corresponding to bit 2 of the entry's signature vector that is if any of the bi-grams AI through AP are set in the bi-gram vector bit 2 of the entry's signature vector will have been set and bit $g_{1,1}$ of the first group of the entry's partitioned signature vector will be set the same process is used to assemble the other bits of the first group, and then the bits of the other six groups thus the first group gives the best feature subset 'bit positions in a signature vector' based on the statistics on a given lexicon the next group gives the second best subset and so on; which is readable as moving a relative position of the entry for the string within the sequence of slots toward the beginning of the sequence of slots) (see col. 8, lines 26-41).

As per claims 11 and 26, Li substantially teaches a method as claimed, further comprises the step of initializing a descriptor for the hash table, said descriptor storing a reference to the hash table and parameters for the hash table (thus, in the lexicon is stored in a bucket associated

with the bucket address equal to the decimal number of the group, which is readable as descriptor storing a reference to the hash table and parameters for the hash table) (see figure 4B, cols. 7 and 8, lines 50-67 and 1-12);

wherein the step of identifying a hash table includes the step of identifying a descriptor indicating the hash table and a prime number (thus, the bucket addresses associated with the unverified string is stored the content and length of this list of pointers will depend on the number of groups selected to be hashed, which is readable as identifying a hash table includes the step of identifying a descriptor indicating the hash table and a prime number) (see col. 9, lines 18-21).

As per claims 12 and 27, Li substantially teaches a method as claimed, wherein the step of initializing a descriptor for the hash table includes the step of initializing a prime number for use in computing a hash value (thus, the partitioning and hashing subroutine of figure 2 is run for each entry of the lexicon the indexing or hashing step 240 is carried out for all the entries of the lexicon when the entire lexicon is processed, many of the individual bucket addresses will be associated with varying numbers of the entries of the lexicon all the foregoing steps are completed in advance of actual matching to any unverified string the group partitioning and the linked list table 40 are fixed for a particular lexicon; which is readable as wherein the step of initializing a descriptor for the hash table includes the step of initializing a prime number for use in computing a hash value) (see col. 8, lines 19-27).

As per claims 13 and 28, in addition to the discussion in claim 6 above, Li teaches wherein the step of initializing a descriptor for the hash table includes the step of initializing a maximum

number of slots for the hash table (where, the threshold may be set to limit the final list to the set number of candidates preferably ten candidates, which is readable as initializing a descriptor for the hash table includes the step of initializing a maximum number of slots for the hash table) (see col. 10, lines 4-10).

As per claims 14 and 29, the limitations of claims 14 and 29 are rejected in the analysis of claim 6 above, and these claims are rejected on that basis.

As per claims 15 and 30, in addition to the discussion in claim 1, Li teaches the steps of said hash table containing sequences of slots for holding respective key values, each of said sequences of slots corresponding to a respective hash value (thus, partitioning the signature vector into groups each having a predetermined number of bits, the bits of the groups are preferably arranged in descending order of frequency of appearance of each bit in the lexicon; which is readable as each of said sequences of slots corresponding to a respective hash value) (see col. 3, lines 5-8);

using said hash value to locate a beginning of the particular sequence of slots that correspond to said hash value (thus, partition the signature vector of a particular lexicon entry the same bits are compared to the bit mask only if the bit is set to '1' in the entry's signature vector is the corresponding bit of the first group set to '1' for example if bit no. 2 'AI through AP' contains the bi-grams having the most frequent occurrence in the lexicon bit $g_{1,1}$ will be given a value corresponding to bit 2 of the entry's signature vector that is if any of the bi-grams AI through AP are set in the bi-gram vector bit 2 of the entry's signature vector will have been set and bit $g_{1,1}$ of

the first group of the entry's partitioned signature vector will be set the same process is used to assemble the other bits of the first group, and then the bits of the other six groups thus the first group gives the best feature subset 'bit positions in a signature vector' based on the statistics on a given lexicon the next group gives the second best subset and so on; which is readable as using said hash value to locate a beginning of the particular sequence of slots that correspond to said hash value) (see col. 8, lines 26-41);

searching the particular sequence of slots for a slot holding a key value matching the key, and if a slot having a key value matching the key is found in the particular sequence of slots, but is not at the beginning of said particular sequence of slots, then moving a relative position of the key value within the particular sequence of slots toward the beginning of the particular sequence of slots (thus, determined whether all the entries of the lexicon have been processed, if not the method returns to step and the next lexicon entry is selected, the partitioning and hashing subroutine is run for each entry of the lexicon, the indexing or hashing is carried out for all the entries of the lexicon when the entire lexicon is processed many of the individual bucket addresses will be associated with varying numbers of the entries of the lexicon, all the forgoing steps are completed in advance of actual matching to any unverified string; which is readable as searching the particular sequence of slots for a slot holding a key value matching the key, and if a slot having a key value matching the key is found in the particular sequence of slots, but is not at the beginning of said particular sequence of slots, then moving a relative position of the key value

within the particular sequence of slots toward the beginning of the particular sequence of slots (see col. 8, lines 16-24).

As per claims 33 and 34, in addition to the discussion in claim 1 above, Li teaches the first lexical container is configured to hold more entries that the second lexical container (thus, a second portion of the lexicon comprising some of the entries of the first portion by directly comparing an encoded representation of the unverified string with encoded representations of the entries of the first portion of the lexicon, which is readable as the first lexical container is configured to hold more entries that the second lexical container) (see col. 4, lines 29-33).

5. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure. Ahuja et al. US Patent Number 5,946,679 relates generally to searching of tables and specifically to route tables given a large field to match against as a search key.

Conclusion

6. Any inquiry concerning this communication from examiner should be directed to Jean Bolte Fleurantin at (703) 308-6718. The examiner can normally be reached on Monday through Friday from 7:30 A.M. to 6:00 P.M.

If any attempt to reach the examiner by telephone is unsuccessful, the examiner's supervisor, Mrs. KIM VU can be reached at **(703) 305-8449.** The FAX phone numbers for the Group 2100 Customer Service Center are: ***After Final*** **(703) 746-7238**, ***Official*** **(703) 746-7239**, and ***Non-Official (703) 746-7240***. NOTE: Documents transmitted by facsimile will be entered as official documents on the file wrapper unless clearly marked "***DRAFT***".

Any inquiry of a general nature or relating to the status of this application or proceeding should be directed to the Group 2100 Customer Service Center receptionist whose telephone numbers are **(703) 306-5631, (703) 306-5632, (703) 306-5633**.

Jean Bolte Fleurantin

May 28, 2002

JBF/

KIM VU
SUPERVISORY PATENT EXAMINER
TECHNOLOGY CENTER 2100